Stockholm School of Economics
Department of Accounting and Financial Management
Bachelor Thesis in Accounting and Financial Management
December 2023

Do Swedish fund managers create value?

A study on the skill of fund managers of actively managed Swedish funds

Charlie Blidén
25249@student.hhs.se

Christoffer Wielbass
25251@student.hhs.se

Abstract

This thesis investigates the performance and skill of fund managers in the Swedish mutual fund industry, challenging the traditional view that actively managed funds underperform the market. Utilizing a novel approach, the study defines skill as 'realized value added', which incorporates gross alpha and assets under management. The research focuses on two main aspects: the overall skill of Swedish fund managers and the comparative performance of environmentally focused (green) and traditional (brown) funds during different market conditions, including financial crises. Key findings reveal that Swedish fund managers demonstrate the ability to generate positive value on average. The study also uncovers significant variations in fund performance during different market conditions, highlighting the dynamic nature of fund management. Additionally, it provides insights into the performance of green funds, which exhibit distinct patterns compared to their brown counterparts, especially during periods of financial instability.

Tutor: Milda Tylaite
Keywords: Mutual Funds, Value Added, Managerial Skill, Gross Alpha, Assets Under Management, Crisis and No-crisis, Green and Brown Funds
Acknowledgments: A special thanks to Milda Tylaite for valuable guidance and feedback

1. Introduction

The reliance on mutual funds as investment vehicles has experienced explosive growth the past decades. As an investor, there are several aspects to consider when deciding to invest in a mutual fund - should it seek to satisfy a certain level of risk exposure? What degree of geographical and industry diversification is adequate? For what time horizon should the capital be invested in the fund? Amidst the increasing prominence of environmental, social, and governance (ESG) considerations, should one invest in green funds that align with sustainable principles or in traditional brown funds in pursuit of financial returns? Above all, there is the unsettled passive vis-a-vis active management debate: should one invest in an index aspiring to match a specific market return, or entrust capital to fund managers who actively allocate it? Choosing the latter implies an expectation of skill (through stock picking and or market timing) in order to generate abnormal returns above the market rate of return (or the would be return of the fund's respective benchmark), or else the investor would be better off investing passively.

The emergence of vast amounts of mutual funds has prompted substantial discourse concerning fund performance which has long been fueled by multifaceted views. The recent debate has culminated in two dominant topics; are managers capable of generating a risk-adjusted abnormal return less associated costs (net alpha), and are they subsequently capable of consistently producing said alpha for an extended duration? An interesting point is whether the abnormal performance is attributed to pure luck or if there exist stock picking and market timing skills, a point that the literature to this day has failed to provide an absolute answer to. The existing literature, which is predominantly conducted and sourced using US data, present dichotomous views. A substantial body of research posits that active fund managers lack skill, resulting in actively managed funds underperforming their passive counterpart net of fees (inter alia: Jensen (1968), Malkiel (1995), and Fama and French (2010)). That is to say, active fund managers are not capable of beating the market. Conversely, other studies claim that a select few fund managers demonstrate exceptional proficiency in the domain of stock selection and market timing. Because of the conflicting results from research in combination with the scarce research on the Swedish market, it becomes interesting to investigate the presence of skill in the Swedish mutual fund industry.

Furthermore, a majority of existing research uses some form of alpha (net, gross, 3-factor, 4-factor et cetera) to measure skill, a method which we believe has limitations. In line with Berk and van Binsbergen (2015) we argue that alpha is not a measure of skill under the standard neoclassical assumptions (rational investors, competitive financial markets, and managers optimize). Berk and Green (2004) argue that if skill is short in supply, the net alpha is determined by competition between investors in equilibrium, and not by the skill of managers. Still, some suggest that gross alpha (average abnormal return before fees) might serve as a better measure of skill. Berk and Van Binsbergen (2015) postulate that an implication of investor competition is that gross alpha can only differentiate managers if, and only if, the size of the funds are identical. This implies that the gross alpha is equal to the fund's fee and thus the fund manager selects the gross alpha. Furthermore, gross alpha is a return measure, not a value measure. Hence, we adopt

a similar definition as Berk and van Binsbergen (2015) of skill as a measure and define it as: *Realized value added = Assets under management * Gross alpha*. A skill proxy should quantify the capital a fund garners from financial markets which this definition does[1]. By this standard, return measures like alpha fall short, as they gauge performance against a benchmark rather than the absolute monetary value obtained.

The contribution of this paper is multifold. Firstly, there's a significant gap in the literature when it comes to managerial skill in the Swedish fund market. The majority of existing literature mainly covers the US mutual fund industry, with a selection of notable work being Jensen (1968), Carhart (1997), Fama and French (2010) as well as Berk and van Binsbergen (2015). Analyzing the Swedish market would allow for comparative analyses with other markets, like the US. Sweden is one of the countries where funds as a form of savings is the most popular in the world, making it an area of interest to study[2]. Furthermore, this could lead to insights into the universality (or lack thereof) of certain fund management practices and outcomes. Moreover, our study covers a more recent and extended timeframe. By examining up-to-date data over a prolonged period, we seek to deepen insights into the performance of Swedish mutual funds during our selected timeframe. Furthermore, we incorporate a more substantial sample size of funds than previous studies focused on the Swedish market to increase the statistical power of our tests. Secondly, we employ Berk and van Binsbergen's (2015) measure of skill, *realized value added*, thus challenging the prevailing methodologies to potentially offer a more comprehensive measure of fund manager skill. Most literature on Swedish mutual fund performance uses some kind of alpha as a proxy for skill (inter alia: Dahlqvist et al (2000). Furthermore, in line with Berk and van Binsbergen (2015) we depart from existing literature by ascertaining making sure our benchmark was marketable and tradable at the time. This ensures that the benchmark represents an actual investment option available to investors during the time period and that all funds in the study are evaluated against a consistent standard. This is especially important when comparing funds across different time periods, as it ensures that shifting market conditions or the introduction of new financial products don't distort the comparisons.

Moreover, we depart from Berk and Van Binsbergen's (2015) work in several ways: rather than focusing on the persistence of skill, we aim to map the different characteristics in value added within different dimensions. One of these is the ESG dimension, where we investigate whether

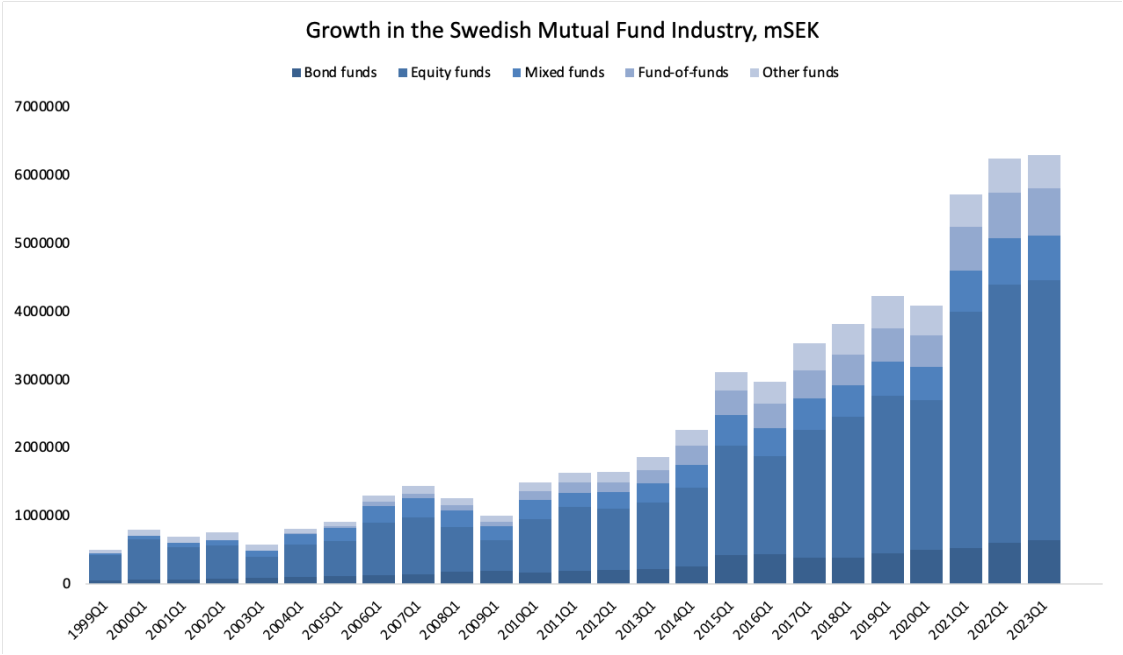
[1] Hypothetically, consider a fund manager who oversees an investment fund. In the early years, the fund had a relatively modest size of 10 million SEK. During this time, the investment strategies result in a gross alpha of 5%, indicating that the fund is performing 5% better than its benchmark. Despite the impressive percentage, the actual value added to the fund is 0.5 million SEK (5% of 10 million SEK). As the manager's reputation grows, so does the fund's assets under management, increasing to 100 million SEK. With the increased scrutiny and challenges of managing a larger fund, the gross alpha drops to 2%. At first glance, the decrease in gross alpha might suggest a loss of skill. However, the value added by the fund is now 2 million SEK (2% of 100 million SEK), which is four times the value added in the fund's earlier days. This example demonstrates that while the gross alpha has decreased, the actual realized value added—and by extension, the fund manager's skill in generating monetary value for the fund—has increased. The growth in fund size amplifies the impact of gross alpha, as even a smaller percentage of outperformance translates to a larger absolute return for investors. Gross alpha alone might suggest that the fund manager's early success was luck, but the realized value added measure offers a more comprehensive and persistent measure of skill compared to traditional measures like alpha, as it captures the tangible monetary value a fund manager extracts over time, reflecting the culmination of skill used to extract money from markets. As aforementioned, investor competition drives net alpha to zero if skill is short in supply and can therefore not be used as an adequate measure of skill.

[2] 80% of Swedish people have fund savings privately and if mandatory savings for the premium pension is included, all Swedish people are fund savers (in Sweden a portion, 2.5% of the salary, goes to the premium pension). (Swedish Investment Fund Association)

ESG("green")-funds outperform their brown counterparts. Moreover, another dimension we explore is whether the value added among fund managers is impacted by varying market conditions. Specifically, we examine periods of financial crises and contrast them with periods of no-crises. Additionally, we merge these two dimensions into one by examining the performance in terms of value added of green and brown funds during times of crisis and non-crisis. Lastly, we conclude with discussing return as a concept in a much broader accounting context and what other returns could be considered by stakeholders in the context of mutual fund investing.

Our study is structured into five different sections. The second section provides an overview of prior research relevant to our thesis. In the third section, we explain our method and sample. The fourth section is dedicated to presenting and analyzing our results. Finally, the fifth section presents our conclusions and discusses its potential implications.

Figure 1: Growth in the Swedish mutual fund industry



This figure reports the growth in the Swedish mutual fund industry ranging from Q1 1999 to Q1 2023, reported in SEK million. Additionally, it displays the distribution of different funds, including bond funds, equity funds, mixed funds, fund-of-funds, and other funds.

2. Literature review

In this section, we will introduce prior literature relevant to our study. We will begin by reviewing global research, including the efficient market hypothesis, Jensen's alpha, Carhart's 4-factor model, and Berk and van Binsbergen's realized value added. We will then continue with literature concerning the performance of green and brown funds and lastly, present a study on the performance of active vis-a-vis passive funds on the Swedish market.

The efficient market hypothesis is a foundational concept in modern financial economics, and as a whole it theorizes that the market is efficient. The implication of efficiency is that the market, over time, cannot be beaten as it incorporates all price-relevant material information, and that any outperformance is temporary and caused by luck (Fama, 1970). In this light, sustaining abnormal returns over significant periods of time would be impossible. Yet, a plethora of investors have been able to beat the market, some prominent examples being Warren Buffett, Peter Lynch and David Swensen. Consequently, the hypothesis has faced scrutiny on both theoretical and empirical grounds; researchers have found market anomalies that capture deviations from predicted asset prices (inter alia: F.M. De Bondt and Thaler (1985), Jegadeesh and Titman (1993), Barberis, Shleifer and Vishny (1998)), indicating that the market is not always efficient.

Michael C. Jensen introduced the topic of outperformance specifically within the mutual fund industry, and was a trailblazer in the field. He introduced "*Jensen's alpha*", a novel *absolute* risk-adjusted performance measure, to assess if fund managers exhibit predictive skills in evaluating future mutual fund performance (Jensen, 1968). His findings indicated that, on average, mutual funds do not demonstrate a significant ability to accurately predict security prices, leading to an inability to consistently beat the market. Today, the predominant model employed for risk adjustment when evaluating fund performance is attributed to Carhart (1997) and is commonly known as "Carhart's four-factor model". This model serves as an extension of the Fama-French three-factor model by introducing an extra factor, known as the momentum factor. The factor aims to capture the one-year momentum anomaly introduced by Jegadeesh and Titman (1993). The models make use of several market anomalies that capture market deviations, these being SMB, HML and the MOM anomaly for the four-factor model. Carhart's research reveals that these four factors almost exclusively account for the persistence in performance observed among mutual funds.

Our analysis primarily builds upon the foundational paper by Berk and van Binsbergen (2015), where they investigate the skill of the US mutual fund industry. With the definition of skill as the value a mutual fund extracts from capital markets, they found that they on average generate approximately $3.2 million per year (inflation-adjusted to USD in 2000). Furthermore, contrary to other literature the authors found large cross-sectional differences in skill persist for as long as 10 years. Additionally, they find that investors are able to recognize skilled fund managers and thus more capital is directed into these funds. Consequently, better funds receive higher aggregate fees; there is a strong positive correlation between current compensation and future performance exists. The cross-sectional distribution of managerial skills is predominantly reflected in the cross-sectional distribution of fund size, not in gross alpha.

Climent and Soriano (2011) report that increased environmental concern among investors has led to the vast emergence of so-called “green funds” as investors take into consideration other aspects than financial return, including environmental, social and governmental (ESG) aspects when selecting their investments. Climent and Soriano (2011) assess the performance of green funds based in the U.S. and compare it to their traditional counterparts - conventional mutual

funds. With a methodology based on the Capital Asset Pricing Model (CAPM), their findings indicate that during the period from 1987 to 2001 green funds underperformed, but from 2001 to 2009, there was no significant difference in performance between them. One of the earliest studies within the field was White (1995) who conducted a comparative analysis in the U.S and Germany of the financial performance between the market and green funds. The findings of the study were that green funds based in the U.S. have a lower return than the market index, whereas in Germany green funds show no significant deviation from the market index. This holds particular relevance to our analysis, as we will explore if the value created by brown and green funds differs throughout the time period.

Adler and Kritzman (2008) assess how socially responsible investing (SRI) impacts the performance of the investment portfolio, positing that SRI implies a trade-off between financial performance and social responsibility. Through Monte Carlo simulation, they approximate that the investment cost of SRI ranges between 0.08% and 2.71% in returns per year, attributing this to restrictions made by green funds. Green funds have self-imposed restrictions for their investments, including restrictions of certain non-ethical sectors such as alcohol, tobacco, weapons, pornography, nuclear weapons, gambling and defense industries. Another noteworthy contribution to this topic comes from Glode (2011), who finds that the increased demand for funds that are actively managed can be ascribed to the ability these managers have to perform better in adverse market conditions rather than in favorable market conditions. Notably, the findings echo those in Nofsinger and Varma (2014), who analyze equity funds based in the U.S. from 2000 to 2011 and investigate if SRI funds mitigate downside risk for investors during time periods of crisis. They find that SRI funds have higher returns than conventional mutual funds during time periods of crisis, but the reduction of downside comes at the expense of underperformance during time periods of no-crisis. In our contribution to this discourse, we introduce an extra dimension through analyzing performance using the absolute measure of value added instead of alpha. This methodology extends the scope of analysis, yielding a more comprehensive perspective on the interplay between financial performance and social responsibility, particularly within the context of financial instability.

In spite of the growing significance of mutual fund investments among both Swedish households and institutional investors, there remains a notable scarcity of research concerning the performance within the Swedish mutual fund industry. One exception is the study conducted by Dahlquist, Engström, and Söderlind (2000), which delved into the relationship between fund performance and various fund characteristics in the Swedish market between 1992-1997. In their analysis, performance was assessed by calculating the alpha derived from a linear regression of fund returns against a set of benchmark assets. The findings of Dahlquist et al. (2000) unveiled instances of outperformance within small equity funds, funds characterized by lower fees, those demonstrating heightened trading activity, and, in select instances, funds with a history of robust past performance. Furthermore, their research hinted at the presence of performance persistence, particularly notable in the context of money market funds, albeit less conspicuous within other fund categories. In addition to using a value-based measure of skill, our contribution to the study

of performance of funds on the Swedish market is that we have a longer time period and thus have a more substantial sample size as well as a more recent data set.

3. Method

In this section, we introduce the methodology used in this study. We begin by introducing various variables before going into the main hypothesis of this paper. We then continue with the choice of benchmarks of which the returns of the funds are compared, and lastly present the sample selection and the sub-hypotheses.

3.1 Variables

Excess return (R^n_{it})

Excess return, R^n_{it}, is defined as the return above the risk-free rate for the *i*th fund in month *t*, where *t* represents one month in the sample period, thus taking values from 1 to 264. *i* represents each individual fund. R^n_{it} consists of two parts: firstly R^B_{it}, which is the fund manager's benchmark, defined as the excess return of the next best alternative investment. The second part is ε_{it}, which represents the residual deviation from the benchmark.

(1) $$R^n_{it} = R^B_{it} + \varepsilon_{it}$$

Net alpha (α^n_i)

Net alpha, α^n_i, is defined as the conditional mean of ε_{it}, which can be estimated by (assuming benchmark return is recorded, an assumption which is relaxed later):

(2) $$\hat{\alpha}_i^n = \frac{1}{T_i}\sum_{t=1}^{T_i}\left(R_{it}^n - R_{it}^B\right) = \frac{1}{T_i}\sum_{t=1}^{T_i}\varepsilon_{it},$$

Gross excess return (R^g_{it})

Gross excess return, R^g_{it}, is defined as the excess return fund *i* delivers ex-ante the management fee, $f_{i,t-1}$, which is charged from time *t-1* to *t*:

(3) $$R^g_{it} \equiv R^n_{it} + f_{i,t-1} = R^B_{it} + \varepsilon_{it} + f_{i,t-1}$$

Gross alpha (α^g_i)

Gross alpha, α^g_i, is defined as the benchmark adjusted gross excess return earned by fund *i* ex-ante the management fee. As such, gross alpha can be estimated by:

$$\hat{\alpha}_i^g = \frac{1}{T_i}\sum_{t=1}^{T_i}\left(R_{it}^g - R_{it}^B\right) = \frac{1}{T_i}\sum_{t=1}^{T_i}\left(f_{i,t-1} + \varepsilon_{it}\right).$$

(4)

Realized value added (V_{it})

To compute the *realized value added* (V_{it}) between times *t-1* and *t*, two factors are taken into account. Firstly, we factor in the gross alpha. Secondly, we consider the assets under management (AUM) - the actual fund size at the end of the previous period. This calculation represents the absolute SEK value of the performance relative to the benchmark. By using equation (3), we derive the following:

(5) $$V_{it} \equiv q_{i,t-1}\left(R^g{}_{it} - R^B{}_{it}\right) = q_{i,t-1} * f_{i,t-1} + q_{i,t-1} * \varepsilon_{it}$$

From equation (5) we can infer that fund managers contribute to the realized value added through two distinct elements: 1) $q_{i,t-1} * f_{i,t-1}$ is the managerial compensation, representing the fees the fund accumulates 2) $q_{i,t-1} * \varepsilon_{it}$ is the value the fund provides to investors, which can be both positive or negative.

Skill (S_i) and estimated value added ($\hat{S}_i$)

Skill is measured as the time series expected value of the realized value added from equation (5):

(6) $$S_i \equiv E[V_{it}]$$

This estimated value added for a fund which exists for T_i periods is calculated by

$$\hat{S}_i = \sum_{t=1}^{T_i}\frac{V_{it}}{T_i}.$$

(7)

Estimating average value added across funds - Ex ante distribution

The mean of the distribution from which value added is drawn can consistently be estimated by:

$$\overline{S} = \frac{1}{N}\sum_{i=1}^{N}\hat{S}_i,$$

(8)

where N = # of mutual funds in data.

Estimating average value added across funds - Ex post distribution

For the mean of surviving funds, the average value added can be estimated by weighting each fund by how many periods it appears in the data:

$$\overline{S}_W = \frac{\sum_{i=1}^{N} T_i\hat{S}_i}{\sum_{i=1}^{N} T_i}.$$

(9)

3.2 Hypothesis

In the following section, we delineate our hypothesis. To empirically investigate the true value fund managers may or may not add we adopt the no-skill hypothesis.

Table 1: The no-skill hypothesis.

$\boldsymbol{H_0}$	No fund manager has skill
$\boldsymbol{H_1}$	Alternative hypothesis

No-skill hypothesis

$$\boldsymbol{H_0^{NS}}: S_i = 0$$
$$\boldsymbol{H_1^{NS}}: S_i \neq 0$$

$\boldsymbol{H_0^{NS}}$ postulates that no individual fund manager has skill. This is the dominant hypothesis within the topic area and was initially coined by Fama (1965, 1970), and was also considered in Fama and French's (2010) paper on skill in the U.S. market:

(10) $$S_i = 0$$

$\boldsymbol{H_1^{NS}}$ is the alternative hypothesis we entertain and posits the existence or nonexistence of managerial skill where fund managers either use their abilities to provide or destroy value. If the measurement of managerial skill proves challenging, it stands to reason that less skillful managers might exploit this uncertainty to their advantage. Therefore, we encounter the following:

(11) $$S_i \neq 0$$

3.3 Choice of benchmarks

To determine whether any given fund creates or extracts value from the market, one must compare the performance to the next best investment opportunity available – the benchmark. Up until now, we have been working under the assumption that we have exact knowledge of the benchmark return. In practicality, it is unknown. Therefore, in this section, we will outline three methodologies we employ to ascertain and define the benchmark.

Benchmark 1 - Vanguard index funds

Vanguard index funds are regarded as the cheapest way to hold a well-diversified portfolio and are therefore used as an alternative investment opportunity set. Berk and Van Binsbergen (2015), who examine the performance of U.S. funds that invest both domestically and internationally, use

the 11 Vanguard index funds visible in Table 2 as their first benchmark. We utilize their approach and use an identical set of funds. The rationale behind this is that Swedish funds, similar to U.S. funds, do not only invest in their home country. Thus, adopting a more comprehensive approach through Vanguard funds covering stocks to a global extent as a benchmark is suitable.

Table 2: This table reports the list of index funds used in the construction of the Vanguard benchmark.

Ticker	Asset Class	Fund Name
VFINX	Large-Cap Blend	S&P 500 Index
VEXMX	Mid-Cap Blend	Extended Market Index
NAESX	Small-Cap Blend	Small-Cap Index
VEURX	International	European Stock Index
VPACX	International	Pacific Stock Index
VVIAX	Large-Cap Value	Value Index
VBINX	Balanced	Balanced Index
VEIEX	International	Emerging Markets Stock Index
VIMSX	Mid-Cap Blend	Mid-Cap Index
VISGX	Small-Cap Growth	Small-Cap Growth Index
VISVX	Small-Cap Value	Small-Cap Value Index

We use β^j_i to denote the weight attributed to the *j*th index fund, where *j* ranges from 1 to 11 representing each Vanguard index fund. β^j_i is calculated from the linear projection of the *i*th active mutual fund onto the set of Vanguard index funds. Fund *i*'s benchmark return at time *t* is then determined through the following equation:

$$R^B_{it} = \sum_{j=1}^{n(t)} \beta^j_i R^j_t, \tag{12}$$

R^j_t represents the excess return earned by investors in the *j*th Vanguard index fund at time *t*, *n(t)* is the number of index funds which Vanguard offers at time *t*.

Benchmark 2 - Fama-French-Carhart (FFC)

The traditional risk-based approach is the Fama-French-Carhart (FFC) factor specification which assumes the risk of a portfolio can be measured using a set of risk factors. The benchmark return is equal to the return a portfolio of equal risk:

$$R^B_{it} = \beta^{mkt}_i * MKT_t + \beta^{sml}_i * SMB_t + \beta^{hml}_i * HML_t + \beta^{umd}_i * UMD_t \tag{13}$$

where MKT_t; SMB_t; HML_t, and UMD_t represent the different four factor portfolios: excess market return, small minus big, high minus low, and up minus down, respectively. The β_is represent corresponding risk exposures of the i'th fund to each factor and are estimated through regression analysis.

Benchmark 3 - Capital Asset Pricing Model (CAPM)

The Capital Asset Pricing Model (CAPM) was developed independently by various economists, including Sharpe (1964) and Lintner (1965). It assumes the risk of a portfolio can be measured by an asset's sensitivity to the volatility of the market, which is the only risk factor used in the model. It can be said to be a simpler version of the FFC-model, however, it is still widely used:

$$R^B_{it} = \beta^{mkt}{}_i * MKT_t \tag{14}$$

where MKT_t represents excess return on the market, $\beta^{mkt}{}_i$ is estimated through regression analysis and represents the corresponding risk exposures of the i'th fund to the excess return on the market.

3.4 Data selection and data sources

Sample selection

The selected sample consists of all actively managed open-ended mutual funds with Sweden as domicile, excluding index funds, bond funds, and money market funds. Our study analyzes a 22-year time period starting 1st of January 2001 to 31st of December 2022. This time frame is motivated by data availability concerning the Vanguard index funds and to capture the most recent market dynamics. We correct for the effect of survivorship bias[3] through including both active and inactive (merged, acquired or liquidated) funds in the sample, as such a bias can lead to misleading persistence in performance rankings (Ross et al., 1992). Furthermore, we only include funds with at least three years of historical data on returns in the sample and use 36 monthly observations as a cutoff point for the returns. The motivation behind this is to produce statistically reliable regression coefficients. Moreover, the 36-month cutoff acts as a filter to exclude funds with very short lifespans. Such funds may not offer a representative view of long-term financial performance, and their inclusion could introduce noise and bias into our analysis. Using Morningstar Direct screening for these criteria we obtain a list of 2016 funds for the time period. The data points for the time period on the open-ended funds extracted consists of fund name, ISIN-code, monthly return, monthly total assets under management (AUM), and annual report net expense ratio as well as Morningstar Sustainability Rating™. All data is reported in SEK.[4]

[3] Survivorship bias entails that funds that have ceased to exist or have been merged into other funds are excluded from performance studies due to their non-existence.

[4] There are cases where fund companies offer their mutual funds for sale in different currencies. In Morningstar Direct, all these are included and the same data point on AUM and expense ratio is reported, but the monthly returns vary due to the currency effects. Additionally, there are cases where differently registered funds from the same fund company have almost identical investment holdings, leading to similar returns in the data. This likely increases the multicollinearity, a issue we discuss under limitations (in section 5.3) Nonetheless, since these different funds are

Data cleaning

From the 2016 funds in the initial sample, 1271 funds in the sample have 36 data points or more on monthly return. Thereafter, only those funds that have data on net expense ratio for at least 1 year are included, leaving 863 funds in the list. In order to get the monthly net expense ratio, the annual net expense ratio is divided by 12. The annual net expense ratio is imputed for the years with missing data due to very few observations of expense ratios. Thereafter, we look at the AUM data point which exists for 825 of the funds. Additionally, some funds have AUM reported for each quarter or year or are for other reasons missing data points for some periods where they have returns. Moreover, the average fund was missing 33 months of AUM (this was determined through observing a quoted return for fund *i* in month *x*, but not an AUM in month *x*). Thus, we use the AUM from the latest known period for that fund and impute it to those time periods and impute all months the fund has a return but no corresponding AUM. Observations whose AUM = 0 are also removed. From this, we obtained a final sample of 825 funds with 36 or more observations on monthly return, monthly AUM, and monthly expense ratio which will be used for the analysis.

Table 3: Sample size of mutual funds

Sample	**Funds in sample**
Initial sample from Morningstar Direct	2016
Excluded due to less than 36 monthly returns	-745
Excluded due to no data on net expense ratio	-408
Excluded due to no data on assets under management	-38
Funds in final sample	825

Benchmark data

The data on monthly returns in SEK for the 11 Vanguard index funds during the time period was obtained from Morningstar Direct. We use the one-month Swedish Treasury bill (SE SSVX) rate as a proxy for the risk-free interest rate which is obtained from the Riksbank[5]. This to properly reflect the risk-free investment opportunity from a Swedish investor's perspective, which is the same for all the benchmarks. Regarding the factor models, we obtain the monthly Fama-French-Carhart factors (excluding the risk-free rate which is obtained from the Riksbank) and the factors for the Global CAPM benchmark from Kenneth R. French's database for "developed countries" which are constructed using data from 23 developed countries in four regions[6]. The "developed

registered with Sweden as domicile and are available for sale in different currencies, they are included in the sample as this study is interested in all types of Swedish active funds, regardless of currency hedges and other currency effects.

[5] The Riksbank is the central bank of Sweden.

[6] The 23 countries are: Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, Great Britain, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, Norway, New Zealand, Portugal, Sweden, Singapore, and the United States.

countries" from Kenneth R. French are the factors with the most global coverage and include all the countries (except Luxemburg) in which Swedish investors' portfolio assets abroad are the highest, including the United States, Great Britain, Finland, Germany, Norway, Japan, Denmark, Ireland and France (SCB, 2019). Therefore, these constitute an appropriate representation for the domain where Swedish fund managers invest in and the rationale behind this is similar to the one behind choosing the global Vanguard index funds. Swedish funds can invest globally and therefore Fama-French-Carhart factors that cover a global field of possible investments are an appropriate proxy for the benchmark portfolio.

Swedish benchmarks for robustness

As a robustness check to the three main benchmarks of this study (Vanguard set of index funds, global Fama-French-Carhart factors, and global Capital Asset Pricing Model), we use a corresponding benchmark using Swedish fund and factor data. We do this to increase credibility and to confirm the results of the study. As a complement to Vanguard, we use a set of Swedish index funds with global coverage of stock holdings. By screening in Morningstar Direct for index funds from all open-end funds with Sweden as domicile and with Sweden as country available for sale, with inception date on or before 1st of January 2001, we make sure the benchmarks were marketable and available for sale during the time period. From this screening we obtain a list of (only surviving) 18 potential funds (see Table 36 in the appendix) and their monthly returns in SEK for the time period. Storebrand has, similar to Vanguard, low fees and effective market exposure and diversification, so we will use Storebrand's 5 index funds that existed during the time period to construct a benchmark portfolio. In terms of the complements to the FFC and CAPM benchmarks, we construct Swedish benchmarks that use the risk factors for FFC and CAPM for the Swedish market, rather than for the global market. These are obtained from the Swedish House of Finance data center; however, the data is only available until the end of 2019. Although this is a limitation, we do not use it as our main benchmark, instead we use it as a sanity check and for robustness to our more global benchmarks.

3.5 Sub-hypotheses

To provide additional insights into the main hypothesis, we also have developed four sub-hypotheses.

Table 4: The four sub-hypotheses of this paper.

	Crisis and no-crisis	**Brown and green funds**	**Green and brown funds during crisis**	**Green and brown funds during no-crisis**

H_0	There is no difference in skill between fund managers in time periods of crisis and no-crisis	There is no difference in skill between fund managers of brown and green funds	There is no difference in skill between fund managers of brown and green funds in time periods of crisis	There is no difference in skill between fund managers of brown and green funds in time periods of no-crisis
H_1	Alternative hypothesis	Alternative hypothesis	Alternative hypothesis	Alternative hypothesis

Periodic differences in estimated value added during crisis and no-crisis

To enhance the depth of our analysis, we adopt a dual-period analysis framework, distinguishing between crisis and non-crisis periods. This bifurcation is motivated by the premise that fund managers' skills are distinctly manifested under varying market conditions. Specifically, the tumultuous periods in the sample of the 2008 financial crisis and the 2020 market crash provide an adequate test ground to assess the adaptability of active fund management strategies under challenging market conditions. Conversely, examining the non-crisis periods allows for an appraisal of fund management performance in more stable and predictable market environments. We define the crises periods as the one-year duration from which we start seeing significant negative cross-sectional returns in our data. Specifically, the crises periods are: 2008-09 to 2009-09, as well as 2020-02 to 2021-02. Consequently, the rest of the sample period comprises the no-crisis data. By contrasting these two distinct temporal landscapes, our analysis aims to offer a more nuanced understanding of the differences in performance between these periods, and by extension, how fund managers adapt in the face of changing market dynamics.

Crisis and no-crisis hypothesis

$$H_0^{CN}: S_{i,C} = S_{i,N}$$
$$H_1^{CN}: S_{i,C} \neq S_{i,N}$$

H_0^C postulates that there is no difference in skill, and by extension value added, between fund managers in time periods of crisis and no-crisis. This to analyze if active fund managers perform equally well during periods of adverse as during favorable market conditions. Therefore, we test the following:

(15) $$S_{i,C} = S_{i,N}$$

H_1^C is the alternative hypothesis and posits that there is a difference in skill between fund managers and therefore value added during these periods:

(16) $$S_{i,C} \neq S_{i,N}$$

Performance of green and brown funds

In this section, we consider alternative measures of return and look into whether there exists a systematic effect on the estimated value added among Swedish funds for the different strategies of green and brown funds.

Morningstar Sustainability Ratings assess the environmental, social, and governance (ESG) performance of companies and funds to help investors make informed decisions aligned with their values. These ratings are represented by a globe icon, with a score ranging from one to five globes, where more globes indicate higher sustainability performance. Morningstar evaluates thousands of entities based on a comprehensive set of ESG criteria, including carbon emissions, board diversity, labor practices, and ethical governance. The ratings aim to provide investors with a quick and easily understandable snapshot of a company's or fund's sustainability efforts. By leveraging these ratings, investors can integrate ESG considerations into their investment strategies, promoting responsible and sustainable financial decisions that align with their ethical and environmental preferences.

We propose an intuitive categorization method for distinguishing between green and brown funds using Morningstar's Sustainability Ratings. Morningstar's ratings use a five-globe system, with three globes representing an “Average” sustainability performance. To effectively categorize funds, we classify those with one or two globes, indicating a low or below-average respectively sustainability practices as "brown funds". Conversely, funds with four or five globes, signifying an above average or high commitment to environmental, social, and governance (ESG) considerations is categorized as "green funds,". This categorization approach simplifies the identification of ESG-aligned investment options, providing a practical framework for investors to make choices aligned with their sustainability goals and preferences.

Data on Morningstar Sustainability Ranking exists for 382 funds out of the 825 funds in the sample. Out of those, those with an “Average” ranking are excluded, leaving 250 funds. Those with an “Below Average” and “Low” ranking are classified as brown funds and add up to 44 funds. Those with an “Above Average” and “High” ranking are classified as green funds and add up to 206 funds. As such, the sample seems to be comprised of predominantly green funds rather than brown funds, which could have an effect on the returns and thereby affecting gross alpha and thus realized value added.

Table 5: Sample size of green and brown mutual funds

Sample	**Funds in sample**

Initial sample	825
Excluded due to missing Morningstar Sustainability Ranking	-443
Excluded due to "Average" ranking	-132
Funds in final sample of green and brown funds	250

Green and brown funds hypothesis

$$\boldsymbol{H_0^{GB}}: S_{i,B} = S_{i,G}$$
$$\boldsymbol{H_1^{GB}}: S_{i,B} \neq S_{i,G}$$

$\boldsymbol{H_0^{GB}}$ postulates that there is no difference in skill, and by extension value added, between fund managers of brown and green funds. In line with White (1995) and Climent and Soriano (2011), we compare green and brown funds to determine if there is a significant difference in the performance. Therefore, we test the following:

(17) $$S_{i,B} = S_{i,G}$$

$\boldsymbol{H_1^{GB}}$ is the alternative hypothesis and posits that difference in skill exists and therefore value added between brown and green funds:

(18) $$S_{i,B} \neq S_{i,G}$$

Inspired by Nofsinger and Varma (2014) we also examine if green funds outperform brown funds during periods of crisis and no-crisis with the two following hypotheses.

The green and brown funds during crisis hypothesis

$$\boldsymbol{H_0^{GBC}}: S_{i,GC} = S_{i,BC}$$
$$\boldsymbol{H_1^{GBC}}: S_{i,GC} \neq S_{i,BC}$$

$\boldsymbol{H_0^{GBC}}$ presumes that there is no difference in skill between fund managers of brown and green funds in time periods of crisis. Therefore, we test the following:

(19) $$S_{i,GC} = S_{i,BC}$$

$\boldsymbol{H_1^{GBC}}$ is the alternative hypothesis and postulates that difference in skill exists and therefore value added between brown and green funds during periods of crisis:

(20) $$S_{i,GC} \neq S_{i,BC}$$

The green and brown funds during no-crisis hypothesis

$$\boldsymbol{H_0^{GBN}}: S_{i,GN} = S_{i,BN}$$
$$\boldsymbol{H_1^{GBN}}: S_{i,GN} \neq S_{i,BN}$$

$\boldsymbol{H_0^{GBN}}$ proposes that there is no difference in skill between fund managers of brown and green funds in time periods of no-crisis. Therefore, we test the following:

(21) $$S_{,GN} = S_{i,BN}$$

$\boldsymbol{H_1^{GBN}}$ is the alternative hypothesis and states that difference in skill exists and therefore value added between brown and green funds during periods of no-crisis:

(22) $$S_{i,GN} \neq S_{i,BN}$$

4. Findings and analysis

In this section, we present the findings and analysis of our study. We begin by reviewing descriptive statistics before conducting the hypotheses testing.

Prior to proceeding, we'd like to include a clarifying note regarding the tables in our study. All figures are presented on a monthly basis and are denominated in SEK, unless explicitly specified otherwise. The term "AUM" denotes assets under management. Additionally, with respect to the benchmarks used, "V" signifies Vanguard, while "FFC" and "4F" correspond to Fama-French-Carhart, and "CAPM" refers to the Capital Asset Pricing Model.

4.1 Descriptive statistics

Table 6: Descriptive statistics for the full sample.

Measure	Mean	Median	P1	P5	P95	P99	SD
AUM (MSEK)	4 236	1 307	8,8	46,2	16 447	36 316	16 273
Expense Ratio	0.001297	0.001250	0.000158	0.000417	0.002217	0.003950	0.000671
Excess Return	0.00483	0.00681	-0.13517	-0.07632	0.07389	0.11977	0.04657
Lifetime	157.425	142.000	39.000	53.000	264.000	264.000	73.501

Gross Alpha (V)	0.00187	0.00115	-0.05888	-0.03311	0.03998	0.06953	0.02338
Net Alpha (V)	0.00057	-0.00005	-0.06048	-0.03463	0.03859	0.06795	0.02338
Gross Alpha (FFC)	0.00337	0.00334	-0.07919	-0.04684	0.05308	0.08518	0.03109
Net Alpha (FFC)	0.00213	0.00215	-0.08055	-0.04812	0.05177	0.08371	0.03108
Gross Alpha (CAPM)	0.00241	0.00259	-0.08267	-0.05017	0.05364	0.09212	0.03271
Net Alpha (CAPM)	0.00117	0.00139	-0.08414	-0.05143	0.05225	0.09074	0.03270

This table reports descriptive statistics for AUM, expense ratio, excess return, lifetime, gross and net alpha.

The average fund holds assets of approximately 4.2 billion SEK in a given month, where the median is approximately 1.3 billion SEK, which is significantly lower than the mean. This indicates a skewed distribution where a small number of funds manage a large portion of the assets, while the majority handle much less. There is a notable large spread in terms of standard deviation of the AUM (16.3 billion SEK), suggesting a vast variability in size between the funds in our sample. This can be highlighted by the 95th percentile of funds managing over approximately 16.4 billion SEK in assets, while the 5th percentile of funds managing just above 46.2 million SEK.

The monthly expense ratio, a measure of the cost of fund management, averages 0.1297% (1.5564% annual expense ratio). The close proximity of the mean and median values suggests a relatively symmetrical distribution of expense ratios among the funds. This implies that investors are likely to encounter funds with similar management cost structures, which is a positive sign for market competitiveness. The top percentiles of funds demand fees from 0.2217% to 0.395% monthly (2.66% to 4.74% annually) whereas the cheapest actively managed funds in our sample charge from 0.0158% to 0.0417% monthly (0.1896% to 0.5004% annually).

Excess return represents the net return in excess of the risk-free rate, and the spread of the excess returns amongst the sample is high with a standard deviation of 4.657%. Excess returns show an average of 0.483%, with a slightly higher median, highlighting a positive tilt in the distribution. This suggests that more than half of the funds are performing above the average. The lifetime of the funds, quoted in months, averages 157 months, equivalent to just over 13 years. This long average lifetime is an indication of fund stability, but it also reflects the broad range in fund longevity, as the wide gap between the mean and various percentiles reveals. As the sample period (2001-01-01 to 2022-12-31) totals 264 months or 22 years, we observe that at least 5% of the funds survive for the entire duration of the sample period.

The mean gross alpha is between 0.187% to 0.337% and the median lies in a slightly broader interval between 0.115% to 0.334% between the benchmarks. The gross alpha varies between -5.888% to -8.267% (1st percentile) and 6.953% to 9.212% (99th percentile), indicating a variability in ability to generate returns between the funds. The mean net alpha is between 0.057% to 0.213% and the median lies between -0.0047% to 0.215%. The net alpha varies between -6.048% to -8.414% (first percentile) and 6.795% to 9.074% (99th percentile). While both gross alphas are positive on average, the net alphas, which account for fees, are smaller, and even negative at the lower percentiles. This indicates that while funds may perform well gross of fees, not all are able to deliver value to investors after costs are taken into account.

4.2 Hypothesis testing

Please note that in the below tables for estimated value added we winsorize the sample at the 1st and 99th percentile to limit extreme values and reduce the effect of outliers on the value added. For all three benchmarks we winsorize 18 estimated value added values.

Table 7: Summary statistics of monthly estimated value added ($\hat{S}_i$)

Metric	**Vanguard**	**FFC**	**CAPM**
Mean	4 221 157	11 143 219	8 006 118
Standard Deviation	11 075 248	19 952 646	16 837 222
T-Statistic	10.95	16.04	13.66
1st Percentile	-10 016 939	-5 084 868	-12 703 060
5th Percentile	-4 179 131	-1 198 121	-3 395 408
10th Percentile	-2 010 168	-444 042	-1 134 747
Median	362 218	2 449 001	1 409 143
90th Percentile	14 705 658	35 605 839	28 033 686
95th Percentile	28 400 824	53 824 255	44 172 269
99th Percentile	62 895 032	102 697 073	91 574 950

This table reports descriptive statistics for estimated value added ($\hat{S}_i$) for a fund which exists for T_i periods using the three benchmarks.

Table 8: Estimated average value added across funds - Ex ante distribution ($\overline{S}$)

$\overline{S}_{Vanguard}$	4 581 743
$\overline{S}_{FFC}$	13 791 843
$\overline{S}_{CAPM}$	10 469 250

This table reports the results for the estimated average value added across funds - ex ante distribution ($\overline{S}$) using the using the three benchmarks.

Table 9: Estimated average value added across funds - Ex post distribution ($\overline{S}_w$)

$\overline{S}_{w,Vanguard}$	5 796 567
$\overline{S}_{w,FFC}$	17 157 748
$\overline{S}_{w,CAPM}$	12 656 509

This table reports the results for the estimated average value added across funds - ex post distribution ($\overline{S}_w$) using the three benchmarks.

Table 7's data on estimated monthly value added using different benchmarks reveals significant insights into fund performance. The estimated average monthly added value for the funds among the three benchmarks all show positive values, ranging from approximately 4.2 million to 11.1 million SEK. This indicates that, on average, the funds are generating value beyond the baseline expectation.

Moreover, when examining the median values for all benchmarks, they are noticeably lower than the respective means, ranging from approximately 0.4 million to 2.4 million SEK. This discrepancy between the median and mean is a common characteristic of data with a positive skew. It implies that while the typical monthly value added is relatively moderate, there are occasional months with exceptionally high values that significantly elevate the average. This phenomenon aligns with the findings of Berk and van Binsbergen (2015), indicating that a majority of the capital is managed by skilled individuals, resulting in an overall positive contribution by mutual funds.

The presence of negative values in the lower percentiles suggests that there also can be substantial variability on the downside. The top funds (>10%) are able to generate positive monthly value added ranging from approximately 14.7 million to 35.6 million SEK, while the worst funds (<10%) extract monthly sums between -0.4 million to -2.0 million SEK or lower from the market between the benchmarks. The standard deviation among the funds relative to the benchmarks varies from 11.1 million to 20.0 million SEK. This wide range indicates substantial variability in the monthly value added figures across different funds. It underscores the fact that fund managers exhibit varying levels of skill, and funds differ significantly in their performance relative to the broader market.

The Vanguard benchmark establishes a mean estimated monthly value added of approximately 4.2 million SEK, with a high t-statistic of 10.95, indicating strong statistical significance. The wide range in percentiles, from a substantial negative value at the 1st percentile to a high positive value at the 99th percentile, suggests considerable variability in fund performance. The median of approximately 0.4 million SEK is relatively modest compared to the mean, indicating that while some funds perform exceptionally well, others do not. Using the FFC benchmark, the mean

estimated monthly value added surges to around 11.1 million SEK, with a t-statistic of 16.04. This higher mean, coupled with an even larger standard deviation, points to greater variability in fund performance under this benchmark. The range in percentiles is even more pronounced than in the Vanguard benchmark, highlighting the diverse outcomes for different funds. The CAPM benchmark presents a mean estimated monthly value added of about 8.0 million SEK and a t-statistic of 13.66, reinforcing the trend of significant value addition by funds. However, like the other benchmarks, the wide range in percentiles and a large standard deviation indicate a high degree of variability in fund performance.

Overall, Table 7 illustrates that fund managers in the Swedish market generate positive value added. However, this value added differs significantly between funds indicating a large distribution of skill levels among the managers. The results suggest that fund performance is highly dependent on the specific strategies and decisions of individual fund managers, but at least as important is the choice of benchmark.

In Table 8 and Table 9, we compute the average monthly value added across funds ($\overline{S}$) and the time-weighted monthly value added ($\overline{S}_w$). Recall that the ex-ante ($\overline{S}$) figure represents a simple average that does not account for how long the funds have been in operation. In contrast, the ex-post ($\overline{S}_w$) figure reflects a weighted average that takes into account the length of time that the funds have been active. The average fund contributes, ex-ante, between approximately 4.5 million SEK to 13.8 million SEK monthly, equating to an annual addition of 54 million SEK to 166 million SEK. When considering the longevity of funds in the calculation, similar to Berk and Van Binsbergen's (2015) findings, our values are higher across all benchmarks. Our results in Table 8 suggest that the average manager adds 5.1 million SEK to 17.2 million SEK monthly. This is in line with our expectations, as it reflects the natural attrition of less skilled funds. These funds tend to cease operations earlier, thus diminishing their negative impact on the weighted average ($\overline{S}_w$).

Hypothesis test

To assess whether the fund managers create value or not, we run a series of t-tests. In order to perform traditional two-tailed hypothesis tests, there are several assumptions that need to be satisfied. These are addressed in the appendix under section 7.2.

Table 10: T-test of the no-skill hypothesis

Benchmak	Estimate	T-Statistic	P_value	DF	Conf_low	Conf_high
Vanguard	4 221 157	10.95	3.9075e-26	824	3 464 302	4 978 013
FFC	11 143 219	16.04	1.3388e-50	824	9 779 704	12 506 735
CAPM	8 006 118	13.66	1.9749e-38	824	6 855 503	9 156 733

This table reports the results from a two-sided t-test with 95% confidence interval for estimated value added ($\hat{S}_i$) using the three benchmarks.

In examining the t-test results for the Vanguard, FFC, and CAPM benchmarks, all three showcase positive estimates of mean estimated monthly value added, with estimates ranging from 4.2 million to 11.1 million SEK. These estimates are statistically significant at even lower than the 1% level, strongly suggesting that the observed mean estimated monthly value added is not a product of random chance. With 95% confidence we observe that the upper limit for the mean estimated value added ranges from approximately 5.0 million to 12.5 million SEK, and the lower limit ranges from 3.5 million to 9.8 million SEK. Complementing this, the T-statistics for each benchmark are notably high (Vanguard at 10.95, FFC at 16.04, and CAPM at 13.66), further reinforcing that the mean value added significantly differs from 0. Given these findings, we find support for the alternative hypothesis, and reject the null hypothesis that no manager has skill. The results suggest that some fund managers in our sample indeed exhibit skill, adding an estimated 4.2 million to 11.1 million SEK monthly to the market.

4.3 Sub-hypotheses

Periodic differences in estimated value added during crisis and no-crisis

Swedish fund managers appear to consistently create value for our sample period. In this section, we adopt a dual-period analysis framework, distinguishing between crisis and non-crisis periods, to potentially identify differences in value added during different market conditions.

Table 11: Estimated value added ($\hat{S}_i$) during crisis

Metric	**Vanguard**	**FFC**	**CAPM**
Mean	16 570 110	-8 454 521	-3 918 506
Standard Deviation	51 792 222	97 738 881	144 536 373
1st Percentile	-60 318 386	-145 615 004	-128 736 002
5th Percentile	-13 668 681	-59 697 419	-57 424 454
10th Percentile	-3 777 608	-32 878 513	-24 218 024
Median	2 498 370	-603 868	158 864
90th Percentile	57 895 495	11 921 715	23 670 875
95th Percentile	106 439 518	24 679 463	41 853 024
99th Percentile	250 592 948	129 212 714	191 032 214

This table reports descriptive statistics for estimated value added ($\hat{S}_i$) during the time periods of crisis for a fund which exists for T_i periods using the three benchmarks.

Table 12: Estimated value added ($\hat{S}_i$) during no-crisis

Metric	**Vanguard**	**FFC**	**CAPM**
Mean	3 264 278	16 232 864	12 009 101
Standard Deviation	14 713 086	8 363 8535	88 209 940
1st Percentile	-17 912 762	-5 309 287	-15 789 347
5th Percentile	-5 962 699	-1 171 718	-3 871 917
10th Percentile	-2 850 277	-395 738	-1 377 009

Median	231 974	3 130 582	1 468 586
90th Percentile	10 525 123	44 235 857	29 234 680
95th Percentile	21 717 216	65 413 747	49 493 401
99th Percentile	56 009 121	121 594 418	109 371 010

This table reports descriptive statistics for estimated value added ($\hat{S}_i$) during the time periods of no-crisis for a fund which exists for T_i periods using the three benchmarks.

Table 13: Paired t-test of the mean difference in estimated value added ($\hat{S}_i$) between time periods of crisis and no-crisis

Benchmark	**Estimate**	**Statistic**	**P_value**	**DF**	**Conf_low**	**Conf_high**
Vanguard, mean diff	12 870 896	6.37	3.2958e-10	730	8 905 593	16 836 199
FFC, mean diff	-26 376 840	-3.93	9.2243e-05	730	-39 546 404	-13 207 276
CAPM, mean diff	-17 124 414	-1.98	4.852e-02	730	-34 137 217	-111 610

This table reports the results from a paired t-test for the mean difference in estimated value added ($\hat{S}_i$) between the time periods of crisis and no-crisis using the three benchmarks.

Table 11 and 12 show prominent differences between the two periods. During crises, funds on average added value when compared to the Vanguard benchmark, yet underperformed relative to the FFC and CAPM benchmarks, as evidenced by negative mean values. This period was characterized by high variability in fund performance, with some funds experiencing severe losses and others achieving significant gains. In more stable, no-crisis periods, funds consistently added value across all benchmarks, indicated by positive mean values, with the FFC benchmark showing the highest average value added. The variability in fund performance was notably less during these no-crisis periods. The data suggests that the volatility of the market during crises leads to a wider dispersion of fund performance outcomes, whereas more stable periods see a general trend of value addition with reduced performance dispersion. The choice of benchmark proves to be a critical factor in evaluating fund performance, especially during turbulent market conditions.

However, despite the apparent stark differences in estimated value added between the periods, we cannot infer if the difference between the two periods represents a statistically significant difference just by looking at the numbers in the table. Therefore, we run a paired t-test[7] (table 13) of the mean difference in estimated value added for each fund between the crisis and no-crisis periods to determine if there is a significant difference. The null hypothesis assumes that the true mean difference between the paired samples is zero. With the results being statistically significant at the 1% level for Vanguard and FFC, and at the 5% level under the CAPM benchmark, we find support for that there is a statistically significant difference in value added

[7] The Paired T-test allows for a direct comparison of each fund's performance in one period to its performance in the other, taking into account the individual variability of each fund and isolating the effect of the market condition (crisis vs. no-crisis).

during periods of crisis vs no-crisis. To gain an understanding of the magnitude and direction of the effect, we can observe the estimates in Table 13. According to the Vanguard benchmark, funds appear to outperform during periods of crisis vs no-crisis with statistical significance at the 1% level by 12.9 million SEK monthly. Thus, Vanguard contradicts both the factor-model benchmarks and according to Table 11 the average fund contributes with almost 17 million SEK monthly in value added during crises, while the FFC and CAPM benchmarks stipulate an average fund contribution of approximately -8.5 million SEK and -3.9 million SEK monthly. This is likely caused by the effect of Vanguard representing the in-facto tradable investment opportunity available to investors, including transaction costs and other related costs, which the two other benchmarks do not take into consideration. Consequently, the benchmark return from the Vanguard benchmark is slightly lower than the other two benchmarks, potentially causing the higher estimated value added. Issues with the Vanguard regression model specification is also a possibility, where our model suffers from substantial multicollinearity (see appendix) which could affect the estimates. Nonetheless, for FFC and CAPM we find that funds underperform by 26.4 million SEK (1% significance) and 17.1 million SEK (5% significance) respectively. In other words, the average fund provides 26.4 million SEK and 17.1 million SEK *less* during periods of crisis than during periods of no-crisis for FFC and CAPM.

Performance of green and brown funds

In this section we compare the performance of green and brown funds through descriptive statistics and a series of Welchs' t-tests to assess if the difference in performance is significant. Please note that in the below tables for green and brown funds that we winsorize the sample at the 1st and 99th percentile to limit extreme values and reduce the effect of outliers on the value added. For all three benchmarks we winsorize 8 estimated value added values.

Table 14: Descriptive statistics for green and brown funds

	Mean	**Median**	**P1**	**P5**	**P95**	**P99**	**SD**
G AUM (MSEK)	5 232.7	3 345.3	47.8	153.3	20 478.3	31 491.0	6 235.0
B AUM (MSEK)	2 991.0	1 357.2	172.3	185.5	12 797.7	15 501.8	4 263.0
G Lifetime	199.73	216.50	54.20	88.75	264.00	264.00	67.43
B Lifetime	162.97	141.00	57.70	71.45	264.00	264.00	78.00
G Expense Ratio	0.001156	0.001161	0.000336	0.000345	0.001900	0.005644	0.000738
B Expense Ratio	0.001101	0.001124	0.000519	0.000535	0.001897	0.002036	0.000396
G Gross Alpha (V)	0.063089	0.001768	-0.157355	-0.082753	0.429225	0.710637	0.149984
B Gross Alpha (V)	0.248617	0.010651	0.000434	0.001049	0.801721	0.907793	0.302078

G Gross Alpha (FFC)	0.177154	0.009558	-0.000836	0.000962	0.600339	0.923943	0.227960
B Gross Alpha (FFC)	0.279277	0.056551	0.001038	0.002591	0.839448	0.990004	0.320037
G Gross Alpha (C)	0.140641	0.007095	-0.084554	0.000089	0.534746	0.857367	0.206459
B Gross Alpha (C)	0.248988	0.019715	0.000706	0.002387	0.759425	0.952643	0.285530
G RVA (V)	5 118 754	863 913	-11 593 400	-5 955 728	28 295 632	65 455 509	13 379 239
B RVA (V)	6 825 961	3 498 838	-4 723 565	-2 557 407	31 271 307	38 235 513	10 794 416
G RVA (FFC)	21 341 967	8 999 476	-2 545 350	-97 621	84 881 501	163 749 928	30 983 085
B RVA (FFC)	15 548 810	6 327 792	-30 182	201 916	69 032 583	81 437 018	22 060 018
G RVA (C)	16 102 384	6 738 335	-6 982 212	-409 192	60 519 821	142 113 264	25 243 113
B RVA (C)	10 273 133	3 472 190	-3 840 952	-601 124	41 212 183	61 367 841	16 164 111

This table reports descriptive statistics on green and brown funds AUM, expense ratio, excess return, lifetime, gross alpha and estimated value added ($\hat{S}_i$). Note that AUM is reported in million SEK and that "G" and "B" denote green and brown funds respectively.

Table 15: Welch's two-sample t-test results on estimated value added ($\hat{S}_i$) for green and brown funds.

	Classification	**Estimate**	**Statistic**	**DF**	**P_Value**	**CI_Lower**	**CI_Upper**
Vanguard	Green	5 118 754	-0.91	74.17	3.66e-01	-5 443 874	2 029 461
Vanguard	Brown	6 825 961					
FFC	Green	21 341 967	1.46	83.75	1.48e-01	-2 091 732	13 678 047
FFC	Brown	15 548 810					
CAPM	Green	16 102 384	1.94	94.13	5.54e-02	-136 947	11 795 450
CAPM	Brown	10 273 133					

This table presents the results from a Welch's two-sample t-test for the mean difference in estimated value added ($\hat{S}_i$) for green and brown funds using the three benchmarks. 95% confidence interval.

Table 16: Welch's two-sample t-test results on gross alpha for green and brown funds

	Classification	Estimate	Statistic	DF	P_Value	CI_Lower	CI_Upper
Vanguard	Green	0.13***	-8.98	56.08	1.94e-12	-0.45	-0.28
Vanguard	Brown	0.49***					
FFC	Green	0.35***	-5.31	59.05	1.71e-06	-0.28	-0.13
FFC	Brown	0.55***					
CAPM	Green	0.28***	-6.16	65.49	5.1e-08	-0.28	-0.15
CAPM	Brown	0.49***					

This table presents the results from a Welch's two-sample t-test for the mean difference in gross alpha (quoted in %) for green and brown funds using the three benchmarks. 95% confidence interval.

Table 17: Welch's two-sample t-test results on AUM for green and brown funds

Classification	Estimate	Statistic	DF	P_Value	CI_Lower	CI_Upper
Green	5 232 709 488***	2.89	87.44	4.86e-03	699 970 629	3 783 428 922
Brown	2 991 009 712***					

This table presents the results from a Welch's two-sample t-test for the mean difference in AUM for green and brown funds using the three benchmarks. 95% confidence interval.

According to Table 14, it appears that the average green fund adds more estimated value compared to the average brown fund according to the FFC and CAPM benchmarks, whereas the opposite is true for the Vanguard benchmark. It is worth noting that in the case of the Vanguard benchmark, the values for brown and green funds are relatively similar when compared to the FFC and CAPM benchmarks. Moreover, the average green fund manages roughly 5.2 billion SEK while the average brown fund manages assets of approximately 2.9 billion SEK. Upon examining the differences in gross alpha, there seems to be a substantial difference between green and brown where brown is consistently higher across all percentiles, means and medians for all three benchmarks. Differences in expense ratios seem marginal between the two, suggesting that the cost of managing these funds is similar, implying that there is no significant difference in the level of effort or cost required to manage green assets compared to brown assets.

However, as was the same for the crisis no-crisis analysis, descriptive statistics do not suffice if we want to determine that there is a statistically significant difference between the two. Thus, we move forward to t-testing. For these tests, we will use Welch's t-test as opposed to a regular t-test[8]. A rejection of the null implies a statistically significant difference. As an additional

[8] The motivation behind using the Welch's T-test is twofold: unequal sample sizes, as well as the assumption of equal variances. A Welch's T-test is robust when dealing with unequal sample sizes, which is the case here with 206 green funds and 44 brown funds. Traditional Student's T-tests assume equal sample sizes, and when the sample sizes are significantly different, it can lead to inaccurate results. Moreover, the Welch's T-test

robustness test, we conduct Wilcoxon tests on estimated value added, assets under management (AUM) and gross alpha to assess whether the difference between green and brown funds are statistically significant (see appendix section 7.1). Similar to the Welch's t-test, if the p-value is significant, the null is rejected, and one finds support that the two sample's means are different from each other.

In Table 15, we examine the differences in value added between green and brown funds. The estimate represents the mean estimated value added for green vs brown respectively. For example, we can see that a green fund on average adds 21.3 million SEK per month, whereas a brown fund adds 15.5 million SEK monthly under FFC. The confidence interval represents the interval for the *difference* in value added between green vs brown funds. As shown, each interval across all benchmarks contains 0. Concerning the Wilcoxon tests, we find inconclusive evidence as well (Table 24). With insignificant p-values for all tests on estimated value added, we cannot reject the null and thus we find support that there is no significant difference in value added between green and brown funds.

However, recall that the estimated value added consists of two parts: AUM and gross alpha. Therefore, we further granulate the analysis to examine these variables as well. In Table 16 we test if the green funds' gross alpha significantly differs from the brown funds' gross alpha. The results are clear; across all benchmarks, brown funds deliver substantially higher excess returns in comparison to green funds. For instance, the average brown fund delivered 0.36% higher gross alpha in comparison to its green counterpart under the Vanguard benchmark. The corresponding values for FFC and CAPM are 0.2% and 0.21%, respectively. These results are all significant at the 1% level. Again, the confidence intervals indicate where the *difference* between gross alpha between green and brown lies at a 95% confidence interval. The Wilcoxon tests support this hypothesis (Table 25).

Table 17 highlights that the difference in AUM is significantly different from zero between green and brown funds. Consequently, we find support that the average green fund has higher AUM than the average brown fund. The difference between the two lies somewhere between 699 million SEK and 3.7 billion SEK at a 95% confidence level. Wilcoxon tests support this (Table 26). In conclusion, we find support for brown funds outperforming green funds in terms of return (gross alpha), whereas green funds seem to attract more investments (AUM).

Green and brown funds during crisis

This section provides an analysis of green versus brown funds with respect to their performance during crisis and no-crisis periods. We examine key financial metrics such as estimated value

---

does not assume equal variances between the two groups being compared (green funds and brown funds). To test whether the samples shared equal variances, we tried using Levene's test. However, this test requires equal sample sizes in R, something we do not have. We assume unequal variances and proceed with Welch's two-sample t-test on estimated value added ($\hat{S}i$) for green and brown funds. An assumption of Welch's t-test is normality. Consistent across all specifications (estimated value added, assets under management (AUM) and gross alpha) we have non-normal distributions (we reject the null of the Shapiro-Wilk test). However, as both green and brown samples are larger than 30, we proceed.

added, gross alpha, and assets under management to understand how these funds compare under different market conditions.

Table 18: Welch's two-sample t-test results on estimated value added ($\hat{S}_i$) for green and brown funds during crisis

	Classification	Estimate	Statistic	DF	P_Value	CI_Lower	CI_Upper
Vanguard	Green	21 953 615	-1.06	55.23	2.96e-01	-30 183 119	9 356 309
Vanguard	Brown	32 367 021					
FFC	Green	-8 361 536	0.66	58.26	5.1e-01	-10 818 208	21 530 453
FFC	Brown	-13 717 658					
CAPM	Green	-548 104	0.43	64.77	6.69e-01	-12 490 965	19 335 692
CAPM	Brown	-3 970 467					

This table presents the results from a Welch's two-sample t-test for the mean difference in estimated value added ($\hat{S}_i$) for green and brown funds during crisis using the three benchmarks. 95% confidence interval.

Table 19: Welch's two-sample t-test results on gross alpha for green and brown funds during crisis

	Classification	Estimate	Statistic	DF	P_Value	CI_Lower	CI_Upper
Vanguard	Green	0.35***	-4.85	49.04	1.27e-05	-0.99	-0.41
Vanguard	Brown	1.06***					
FFC	Green	-0.09	-0.15	48.14	8.79e-01	-0.34	0.29
FFC	Brown	-0.06					
CAPM	Green	0.09	-1.2	49.02	2.36e-01	-0.6	0.15
CAPM	Brown	0.31					

This table presents the results from a Welch's two-sample t-test for the mean difference in gross alpha (quoted in %) for green and brown funds during crisis using the three benchmarks. 95% confidence interval.

In terms of estimated value added ($\hat{S}$), during crisis periods, neither green nor brown funds demonstrate a consistent outperformance based on the results in Table 18. The estimates and associated t-statistics do not reveal a significant advantage for either type. Thus, we reject the null and find no statistically significant difference between value added during periods of crises. If we turn to Table 19, the analysis reveals that brown funds have a higher estimate of gross alpha compared to green funds during crisis periods. However, only under the Vanguard benchmark is this effect significant, at the 1% level at that. For FFC and CAPM, we can't conclude that the difference is significantly different from 0 and therefore fail to reject the null. The Wilcoxon tests in Table 27 and 28 reinforce these results.

Green and brown funds during no-crisis

Table 20: Welch's two-sample t-test results on estimated value added ($\hat{S}_i$) for green and brown funds during no-crisis

	Classification	Estimate	Statistic	DF	P_Value	CI_Lower	CI_Upper
Vanguard	Green	3 168 491	0.1	57.3	9.17e-01	-4 220 127	4 686 113
Vanguard	Brown	2 935 498					
FFC	Green	25 187 587	1.09	77.34	2.77e-01	-4 300 188	14 807 462
FFC	Brown	19 933 950					
CAPM	Green	18 032 724	1.51	78.55	1.35e-01	-1 762 023	12 854 607
CAPM	Brown	12 486 432					

This table presents the results from a Welch's two-sample t-test for the mean difference in estimated value added ($\hat{S}_i$) for green and brown funds during no-crisis using the three benchmarks. 95% confidence interval

Table 21: Welch's two-sample t-test results on gross alpha for green and brown funds during no-crisis

	Classification	Estimate	Statistic	DF	P_Value	CI_Lower	CI_Upper
Vanguard	Green	0.10***	-7.84	54.34	1.71e-10	-0.41	-0.24
Vanguard	Brown	0.43***					
FFC	Green	0.40***	-6.4	61.2	2.42e-08	-0.32	-0.17
FFC	Brown	0.65***					
CAPM	Green	0.29***	-5.53	61.31	7.1e-07	-0.31	-0.15
CAPM	Brown	0.52***					

This table presents the results from a Welch's two-sample t-test for the mean difference in gross alpha (quoted in %) for green and brown funds during no-crisis using the three benchmarks. 95% confidence interval

Table 22: Welch's two-sample t-test results on AUM for green and brown funds during crisis and no-crisis

Crisis	Estimate	Statistic	DF	P_Value	CI_Lower	CI_Upper
Green	6 043 929 438***	2.88	85.22	5.01e-03	798 062 068	4 351 840 713
Brown	3 468 978 048***					
No-Crisis	**Estimate**	**Statistic**	**DF**	**P_Value**	**CI_Lower**	**CI_Upper**
Green	5 138 478 224***	2.88	87.69	5.07e-03	678 730 622	3 718 110 081
Brown	2 940 057 872***					

This table presents the results from a Welch's two-sample t-test for the mean difference in AUM for green and brown funds during no-crisis using the three benchmarks. 95% confidence interval

Similarly, as in periods of crisis, in no-crisis periods as shown in Table 20, the t-test results do not strongly favor one over the other when it comes to value added, with the estimates being quite close and the p-values indicating no significant differences. Thus, we cannot reject the null and find no support of significant differences in estimated value added between green and brown in stable market conditions. Concerning gross alpha, Table 21 shows that brown funds still maintain a higher estimate of gross alpha across all benchmarks, implying that brown funds may also outperform green funds when the market is stable. The results for gross alpha are all significant at the 1% level, indicating that there is a substantial difference in gross alpha between the two during periods of no-crisis. Across all benchmarks, the average brown fund delivers excess returns that are 0.25%-0.33% higher than the average green fund, depending on which benchmark.

Finally, we examined potential differences in AUM between green and brown funds during periods of crisis and no-crisis. Regardless of market conditions, we find support for the average green fund managing more assets than the average brown fund, with the *difference* lying somewhere between 798 million SEK - 4.3 billion SEK during crises, and somewhere between 678 million SEK - 3.7 billion SEK during no-crisis at a 95% confidence level. All aforementioned results are supported by the Wilcoxon tests (Table 29, 30 and 31).

4.4 Summary of results

To summarize, our results support a rejection of the no-skill hypothesis at the 1% significance level; there is skill in the Swedish mutual fund industry. These results are unanimous across our three benchmarks. The average fund contributes from 4.2 million SEK - 11.1 million SEK per month, depending on which benchmark. In addition, the estimated average weighted value added across funds ($\overline{S}_w$) ranges from 5.7 million SEK to 17.1 million SEK. As this value is above 0, the results suggest that the average fund manager is skilled.

Moreover, our two factor-model benchmarks show statistically significant results that periodic differences between times of crisis and no-crisis exist and that the average fund provides less (negative) value added during periods of crisis than during periods of no-crisis. However, the Vanguard benchmark, also showing statistically significant results, shows contradictory results. Nonetheless, this is likely caused by lower benchmark return and multicollinearity issues for the Vanguard benchmark. Thus, we reject the null under FFC and CAPM, but not under Vanguard – results are inconclusive.

Furthermore, our results show that differences between brown and green funds in estimated value added is not significantly different from zero. Therefore, our results do not support a rejection of

the hypothesis about differences in estimated value added of brown and green funds. However, when we separate realized value into the two components of its formula (AUM and gross alpha), we find statistically significant results that brown funds outperform green funds in terms of return (gross alpha), whereas green funds attract more investments (AUM).

The difference in skill between fund managers of brown and green funds in periods of crisis is not significant. Hence, we fail to reject the null hypothesis that there is a no difference between value added. Nonetheless, for the Vanguard benchmark, the gross alpha of brown funds is statistically significant. Similar to periods of crisis, we find no significant difference in skill between mangers of brown and green funds in times of no-crisis. Although the skill does not differ significantly, gross alpha is significantly higher for brown funds than for their green counterparts in stable market conditions at the 1% significance level.

Table 23: The results implication on the hypotheses.

Hypothesis	**Description of H_0**	**Conclusion**
No-skill (H_0^{NS})	No fund manager has skill	Rejected
Crisis and no-crisis **(H_0^{CN})**	There is no difference in skill among fund managers in time periods of crisis and no-crisis	Rejected but inconclusive
Brown and green funds **(H_0^{GB})**	There is no difference in skill between fund managers of brown and green funds	Supported
Green and brown funds during crisis **(H_0^{GBC})**	There is no difference in skill between fund managers of brown and green funds in time periods of crisis	Supported
Green and brown funds during no-crisis **(H_0^{GBN})**	There is no difference in skill between fund managers of brown and green funds in time periods of no-crisis	Supported

5. Discussion and conclusions

In this section, we discuss the findings of our study and explore its potential implications. Additionally, we present our conclusions from the findings and discuss limitations and avenues for further research pertaining to our study.

5.1 Discussion of results

No-skill hypothesis

Our results are in line with Berk and van Binsbergen (2015) whose result is 0.27 million USD per month, however, there are some differences. Firstly, the results are not directly comparable to Berk and van Binsbergen (2015) as we use different currencies for the result.
Berk and van Binsbergen (2015) use USD, whereas our study uses SEK. Throughout both the time period of our study and Berk and van Binsbergen's study (2015), there are currency effects which cause differences in the results, making the results not directly comparable.
Furthermore, Berk and van Binsbergen (2015) adjust the assets under management numbers for inflation to USD in 2000. We do not do this inflation adjustment and recognize that it is a limitation in terms of comparability between the papers. However, it is not of significance to the research question to determine the exact real money value, but rather to analyze the differences in the performance of active funds compared to passive funds. Nonetheless, in order to obtain more comparable results over different time periods and with prior literature/research, for future research it would be beneficial to adjust the realized value added for inflation. Lastly, we study different markets and different time periods. Berk and van Binsbergen (2015) studies the U.S. fund market during the time period January 1962 to March 2011, whereas we study the Swedish fund market between January 2001 to December 2022. Differences in both exposure towards different stock markets (in a relative comparison between the two: US funds are more exposed to the US markets whereas Swedish funds are more exposed to the Swedish market) and in performance throughout the time periods also impacts the results. Furthermore, we add a layer to Dahlquist et al's (2000) analysis on fund performance in Sweden, by introducing a new measure of skill and finding that there are skilled Swedish fund managers according to value created.

Periodic differences in estimated value added during crisis and no-crisis

Intuitively, it might appear reasonable that funds underperform during periods of crisis. The logic is that the whole market declines in value due to systematic shocks that are undiversifiable, and thus the funds' returns should, aggregately, also decline. However, one must acknowledge that each fund's benchmark is of adaptable nature, meaning that they incorporate the decline of the market and the estimated value added measure should subsequently represent a fair measure of added value regardless of market conditions. Hence, it's not obvious why funds underperform during periods of crisis which the FFC and CAPM inclines. The findings of Glode (2011) suggest that managers should perform better in periods of crisis compared to periods of no-crisis. Our findings according to the FFC and CAPM benchmark contradict this and show that the funds perform worse with negative value added with statistical significance during periods of crisis. However, the reverse is true for the Vanguard benchmark.

Performance of green and brown funds

In line with White (1995) and Climent and Soriano (2011) we find that green funds underperformed their brown counterparts in terms of gross alpha. Adler and Kritzman's (2008) Monte Carlo simulation that approximates the investment cost of SRI to range between 0.08% and 2.71% in returns per year, also is in line with our results. Nonetheless, green funds add more

value than brown funds because they attract more investment (AUM). The results warrant the question: why do people invest in green funds when they underperform brown funds (in terms of gross alpha)? Are investors in green funds irrational by not maximizing the utility function in the form of financial return? Theories based on sustainable and behavioral finance offer an explanation: investors may have a utility function which consists of not only financial performance, but also social and environmental performance. In this context, Fernandez and Matallin (2008) present a Financial and Social Performance (FSP) measure and if the financial performance of green funds consistently underperforms conventional funds, we could draw conclusions regarding a presence of a financial trade-off between financial and social performance for green funds.

Performance of green and brown funds during crisis and no-crisis

The findings of Nofsinger and Varma (2014) suggest that green funds should perform better than brown funds during time periods of crisis but underperform during time periods of no-crisis. Our findings suggest that during periods of crisis the difference between the performance of green and brown funds is not statistically significant. However, our results depart from Nofsinger and Varma (2014) according to the Vanguard benchmark, where we found that brown funds outperform green funds with statistical significance. Moreover, similar to periods of crisis, our findings on periods of no-crisis suggest that the difference between the performance of green and brown funds is not statistically significant. Nonetheless, in terms of gross alpha, brown funds outperform green funds for all benchmarks with statistical significance, supporting Nofsinger and Varma's (2014) findings. Furthermore, across both time periods, we find that the AUM for green funds is higher than for brown funds with statistical significance.

Method and data discussion

In terms of the reliability of the data, for our main benchmark we have gathered data from Morningstar Direct and the Kenneth R. French data library. These are widely used databases among researchers and professionals for retrieving data on returns and the risk factors. For example, Ibikunle and Steffen's (2017) study on the performance of European green mutual funds uses the identical factors from the same database and Ibert et al's (2017) study on the compensation of fund managers in Sweden also use Morningstar Direct as their primary database. These are a few among many examples and signify the reliability of the sources, and therefore we posit that the data upon which this study is based on is reliable. In addition, we also perform robustness checks on the benchmarks by constructing an alternative benchmark to each benchmark (see appendix 7.4). We find statistically significant results that are in line with those of the main benchmarks.

Moreover, we have followed the computation of value added from Berk and van Binsbergen (2015), which also has been followed by several other researchers including Ibert et al (2017). However, we did a minor modification by dropping funds with less than 3 years of data instead of 2 years as Berk and van Binsbergen (2015). This modification was applied to increase the statistical significance of the tests conducted. This change could potentially impact the results,

which could reduce the validity of the study. However, we argue that this adjustment is minor and has a positive impact on the statistical significance of the study. Furthermore, in terms of hypothesis, we added several in order to provide additional insights of whether the value added varies during differing market conditions and for funds with different characteristics.

5.2 Summary of contributions / Conclusion

This study analyzes the managers of active funds in the Swedish mutual fund industry, a domain previously underexplored in academic literature. By introducing 'realized value added' as a measure of skill to the Swedish market, this study has not only contributed to the existing body of knowledge but has also challenged conventional beliefs about the performance of actively managed funds. Our findings reveal that Swedish fund managers, on average, are capable of generating positive value. This is a significant revelation, suggesting that the skill of fund managers plays a crucial role in the performance of actively managed funds.

The study highlighted the influence of market conditions, particularly financial crises, on fund performance. It was observed that fund managers' ability to add value varied significantly during different market conditions, underscoring the dynamic nature of fund management. The comparative analysis between green (ESG-focused) and brown (non-ESG) funds provided novel insights. Green funds exhibited a distinct performance pattern, especially during periods of financial instability, indicating the growing importance of ESG criteria in investment decisions.

The findings offer a new perspective for investors, emphasizing the need to consider a fund manager's skill and the impact of market conditions when making investment decisions. The performance of green funds also suggests that ESG factors are becoming increasingly relevant in the investment landscape. This study highlights the importance of skill and adaptability in active fund management. The thesis contributes to the broader debate on active versus passive fund management. It underscores the need for more nuanced and comprehensive measures of fund performance, beyond traditional financial metrics.

In conclusion, this thesis not only sheds light on the skill of Swedish fund managers but also opens up new avenues for research in the field of mutual fund performance. It calls for a re-evaluation of how we measure and perceive the success of actively managed funds, advocating for a more holistic approach that takes into account various market conditions and evolving investment criteria, such as ESG factors.

5.3 Limitations and suggestions for future research

Although our sample does not cover all the funds that existed during the time period, the data set is large enough to capture skill and to provide insights into the research question. But despite the inclusion of 825 funds and 264 months in our panel, the most prominent limitation of our study is the possibility that the sample size is not sufficiently large to guarantee a t-distributed t-statistic.

There are two indications suggesting that the t-statistics obtained in our hypothesis testing is overstated. Firstly, it is probable that there is a correlation in value added across funds. Secondly, the distribution of value added exhibits excess kurtosis. From the robustness checks conducted (in the appendix 7.2), we confirm that these issues are present. This is a common challenge in cross-sectional regression, stemming from notable variations in both assets under management and monthly returns across the funds in the sample. Because of this, we suggest that future research should take this into consideration by adopting tests resilient to these characteristics, or expanding the sample size. Furthermore, another limitation is the imbalance in sample size between green and brown funds. This likely has an effect on the descriptive statistics as well as t-tests, where the results are expected to be less stable. Furthermore, a more comprehensive sample size would likely lead to more credible results. A limitation of risk-factor models such as the CAPM and FFC is that they may omit relevant risk factors that explain return. Therefore, for future research, it would be interesting to potentially use additional benchmarks and more comprehensive benchmarks in order to examine the returns against other factors and to increase robustness. This could potentially enhance understanding and provide a more thorough explanation of the achieved returns.

We have defined Swedish funds as mutual funds that have Sweden as their domicile. A fund's domicile is where the fund is registered and does not have to do with where its assets are placed or where the investment decisions are made. It bears another type of significance however, as the domicile dictates the legal framework and regulatory regimes the fund is subject to. It is important to acknowledge that many funds that are managed by Swedish citizens and are available for sale in Sweden have their domicile in other locations, typically Luxembourg or Ireland due to their favorable tax-scheme (Swedish Investment Fund Association, 2022). Since this paper only includes funds with Sweden as domicile, these particular funds are not included in the data set. Therefore, a limitation to this study is the risk of excluding some Swedish fund managers and funds that are targeted towards Swedish investors. Nonetheless, worth reiterating is the aim of this paper: to determine whether Swedish fund managers have skill. Thus, Swedish fund managers who manage funds not domiciled in Sweden, yet have their funds available for sale in Sweden, are encompassed by this research question. As a result, they could be of interest for this study, but they fall outside the scope of the selected sample due to their non-Swedish domicile. Handling funds with international domiciles might also present challenges in terms of data consistency, as regulations and reporting requirements can vary by jurisdiction. Nonetheless, for future research, it would be possible to work more with the data to include Swedish fund managers of funds that are not domiciled in Sweden.

It would also be interesting to study the same phenomena in other geographical areas than Sweden and the U.S. to discover if the results are in line with ours and Berk and van Binsbergen (2015) or if they find contrasting results. Furthermore, it would also be interesting to study other time periods and contrast them with results from others, to see if the skill/value added has changed throughout history. To provide even more nuanced results about the performance of funds, it would also be interesting to investigate whether differences in AUM size impact

performance between funds. In addition, our study has focused on the analysis of value added during varying market conditions (crisis vs no-crisis) and fund characteristics (green vs brown). However, it would be interesting to analyze the skill by investigating the persistence of value added and returns among fund managers. Along this line, more often than not, many Swedish funds boast about beating their benchmark. However, our study shows that many funds in the sample still produce a negative value added. Therefore, for future research, it would be interesting to study whether funds claiming to beat index, actually beat the index and on extension, if they compare their returns to an appropriate benchmark or sporadically change the benchmark index they compare themselves to. Thus, it would be interesting to understand how, when and why funds change their benchmark index.

6. References

Adler, T., & Kritzman, M. (2008). The cost of socially responsible investing. *Journal of Portfolio Management*, 35 (1), 52-56.

Andelar i investeringsfonder efter fondtyp och sektor. (2023). SCB. https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__FM__FM0403__FM0403A/F813FondKv/

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. (1998). "A Model of Investor Sentiment." *Journal of Financial Economics, 49* (3), 307-343.

Behov av nya fondstrukturer i Sverige. (2022) Swedish Investment Fund Association (Fondbolagen). https://www.fondbolagen.se/globalassets/regelverk/framstallningar/behov-associationsrattsliga-fonder-i-sverige.pdf

Berk, J. B., and Green, R. C. (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy, 112*(6), 1269-1295.

Berk, J. B, and van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics, 118*(1), 1-20.

Brown, S. J., Goetzmann, W., Ibbotson, R. G., & Ross, S. A. (1992). Survivorship Bias in Performance Studies. *The Review of Financial Studies*, *5*(4), 553–580.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57-82.

Climent, F., & Soriano, P. (2011). Green and good? The investment performance of US environmental mutual funds. *Journal of Business Ethics*, 103 (2), 275-287.

Dahlquist, M., Engström, S., & Söderlind, P. (2000). Performance and Characteristics of Swedish Mutual Funds. *The Journal of Financial and Quantitative Analysis*, *35*(3), 409–423.

Description of Fama/French 3 Factors for Developed Markets. Kenneth R. French. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_3developed.html

Description of Momentum Factors for Developed Markets. Kenneth R. French. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_developed_mom.html

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383-417.

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, *38*(1), 34–105.

Fama, E. F., and French, K. R. (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns. *The Journal of Finance*, *65*: 1915-1947.

Fama French Factors. Swedish House of Finance. https://www.hhs.se/en/houseoffinance/data-center/fama-french-factors/

Fernandez-Izquierdo, A., Matallin-Saez, J.C. (2008) Performance of Ethical Mutual Funds in Spain: Sacrifice or Premium? *J Bus Ethics* 81, 247–260.

Glode, V. (2011). Why mutual funds "underperform". *Journal of Financial Economics*, 99 (3), 546-559.

Ibert, M., Kaniel, R., Van Nieuwerburgh, S., & Vestman, R. (2018). Are Mutual Fund Managers Paid for Investment Skill? *The Review of Financial Studies*, *31*(2), 715–772.

Ibikunle, G., & Steffen, T. (2017). European Green Mutual Fund Performance: A Comparative Analysis with their Conventional and Black Peers. *Journal of Business Ethics*, *145*(2), 337–355.
Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, *48*(1), 65-91.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance, 23*: 389-416.

Lintner, John. (1965). "Security Prices, Risk, and Maximal Gains From Diversification." *The Journal of Finance* 20.4: 587–615.

Malkiel, B. (1995). "Returns from investing in equity mutual funds 1971 to 1991". *Journal of Finance* 50, 549–572.

Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. J*ournal of Banking & Finance,* 48, 180-193.

Sharpe, William F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of risk." *The Journal of Finance* 19.3: 425–442.

Så blev Sverige världsledande på fondsparande. Swedish Investment Fund Association (Fondbolagen). https://www.fondbolagen.se/fakta_index/sa-blev-sverige-varldsledande-pa-fondsparande/

Sök räntor och valutakurser. Sveriges riksbank. https://www.riksbank.se/sv/statistik/rantor-och-valutakurser/sok-rantor-och-valutakurser/

USA populärt bland svenska placerare. (2019). SCB. https://www.scb.se/hitta-statistik/statistik-efter-amne/betalningsbalans-och-utlandsstallning/betalningsbalans-och-

utlandsstallning/betalningsbalansen-bop/produktrelaterat/Fordjupad-information/usa-populart-bland-svenska-placerare/

Werner F. M. De Bondt, & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, *40*(3), 793–805.

White, M.A. (1995). The performance of environmental mutual funds in the United States and Germany: Is there economic hope for green investors. *Research in Corporate Social Performance and Policy*, 1, 323-344.

7. Appendix

7.1 Wilcoxon tests

In this section we perform Wilcoxon tests to determine whether the Welch's two-sample t-test are statistically significant.

Table 24: Wilcoxon test results on estimated value added ($\hat{S}_i$) for green and brown funds

Wilcoxon Test	**W-Statistic**	**P_Value**
Vanguard	3 663	4.61e-02**
FFC	5 098	1.94e-01
CAPM	5 287	8.31e-02*

This table presents the results from a Wilcoxon test for the mean difference in estimated value added ($\hat{S}_i$) for green and brown funds using the three benchmarks.

Table 25: Wilcoxon test results on gross alpha for green and brown funds

Wilcoxon test	**W-Statistic**	**P_Value**
Vanguard	1 226	3.16e-14***
FFC	2 282	2.39e-07***
CAPM	2 101	2.38e-08***

This table presents the results from a Wilcoxon test for the mean difference in gross alpha for green and brown funds using the three benchmarks.

Table 26: Wilcoxon test results on AUM for green and brown funds

Wilcoxon Test	**Statistic**	**P_Value**
AUM	5 962	0.001***

This table presents the results from a Wilcoxon test for the mean difference in AUM for green and brown funds using the three benchmarks.

Table 27: Wilcoxon test results on estimated value added ($\hat{S}_i$) for green and brown funds during crisis

Wilcoxon Test	**W-Statistic**	**P_Value**
Vanguard	3 975	2.01e-01
FFC	4 020	2.4e-01
CAPM	4 285	5.71e-01

This table presents the results from a Wilcoxon test for the mean difference in estimated value added ($\hat{S}_i$) for green and brown during crisis funds using the three benchmarks.

Table 28: Wilcoxon test results on gross alpha for green and brown funds during crisis

Wilcoxon Test	**W-Statistic**	**P_Value**
Vanguard	2 195	8.04e-08
FFC	4 059	2.78e-01
CAPM	3 924	1.63e-01

This table presents the results from a Wilcoxon test for the mean difference in gross alpha for green and brown funds during crisis using the three benchmarks.

Table 29: Wilcoxon test results on estimated value added ($\hat{S}_i$) for green and brown funds during no-crisis

Wilcoxon Test	**W-Statistic**	**P_Value**
Vanguard	4 131	3.58e-01
FFC	5 006	2.77e-01
CAPM	5 177	1.39e-01

This table presents the results from a Wilcoxon test for the mean difference in estimated value added ($\hat{S}_i$) for green and brown during no-crisis funds using the three benchmarks.

Table 30: Wilcoxon test results on gross alpha for green and brown funds during no-crisis

Wilcoxon Test	**W-Statistic**	**P_Value**
Vanguard	1 324	1.75e-13
FFC	1 888	1.27e-09
CAPM	2 135	3.71e-08

This table presents the results from a Wilcoxon test for the mean difference in gross alpha for green and brown during no-crisis funds using the three benchmarks benchmark.

Table 31: Wilcoxon test results on AUM for green and brown funds during crisis and no-crisis

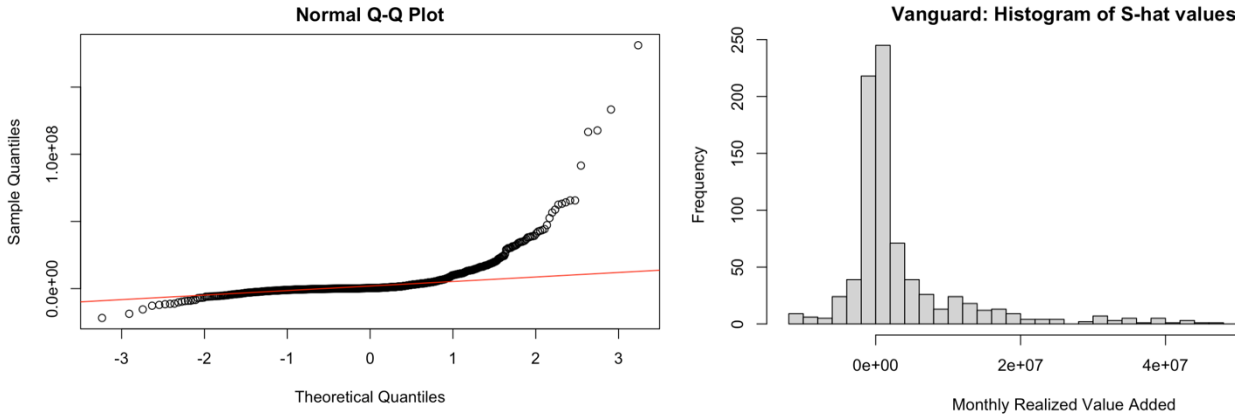
Wilcoxon Test	**W-Statistic**	**P_Value**
Crisis	5 893	1.78e-03
No-Crisis	5 933	1.3e-03

This table presents the results from a Wilcoxon test for the mean difference in AUM for green and brown during crisis and no-crisis funds using the three benchmarks.

7.2 Robustness check for t-tests

Although T-tests are relatively robust to deviations from the assumptions, we perform a test on the assumption of approximately normality through a QQ plot and histogram to satisfy the assumptions to perform a two-tailed hypothesis test.

Figure 2: Example of normal Q-Q plot and histogram of $\hat{S}_i$



The figures above show one of many examples where the assumption of approximately normality is not upheld. Furthermore, as can be seen from the histogram, there is excess kurtosis.

Nonetheless, with reference to the central limit theorem, we can still perform t-tests as long as we fulfill the following assumptions:

- Randomization condition
- Sample independence
- Sufficiently large sample size (n≥30)

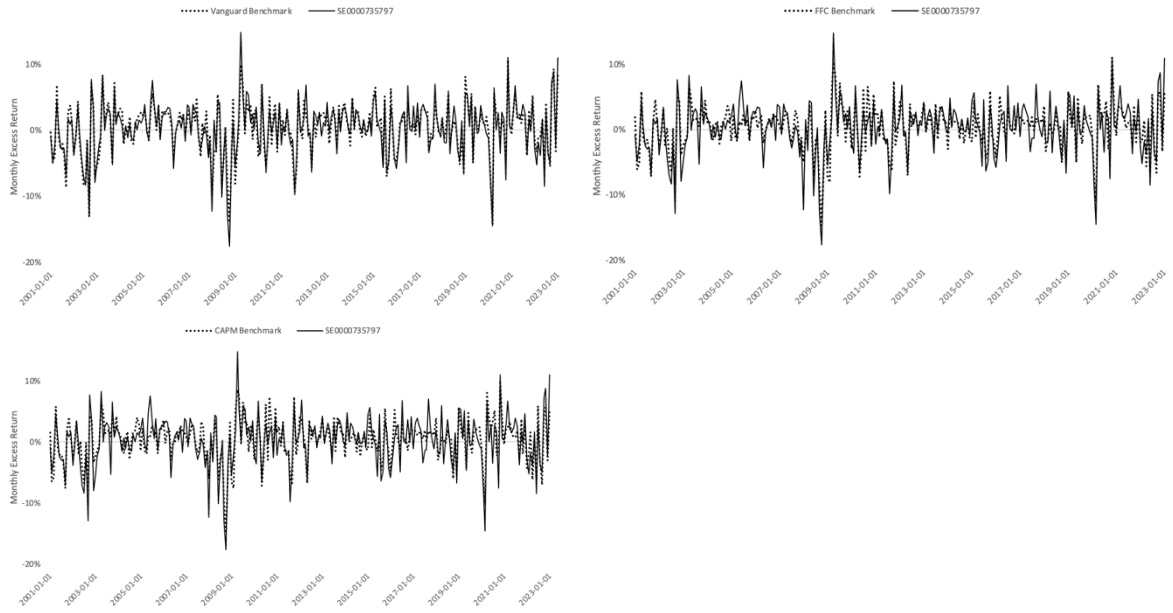
7.3 Robustness check for linear regression models

To investigate the credibility and robustness of our linear regression models we conducted the following diagnostic tests and robustness checks.

Table 32: Tests performed to test the assumptions for linear regression models

Assumption	**Test**
Linear relationship	Plot the observed values versus the predicted values
Normality	Shapiro-Wilk test
Homoscedasticity	Breusch-Pagan test
Auto-correlation	Durbin-Watson test
Multicollinearity	Variation Inflation Factor (VIF) analysis

Figure 3: Examples of linear projection from the three benchmarks.



The figures above compare the excess return of Aktie-Ansvar Europa (ISIN: SE0000735797) with its linear benchmark of Vanguard index funds, FFC, and CAPM. Aktie-Ansvar Europa exists for the whole time period and the benchmarks mimic the excess return of Aktie-Ansvar Europa each month. Linearity holds as we can see that the points are symmetrically distributed around a diagonal line.

Table 33: Summary statistics of robustness checks for the Vanguard benchmark

Test	Min	Q1	Median	Mean	Q3	Max
Shapiro-Wilk	0.00000	0.03810	0.19273	0.30461	0.39517	0.99300
Breusch-Pagan	0.00000	0.11549	0.28293	0.35751	0.46409	0.99501
Durbin-Watson	0.00141	0.73123	0.87291	0.74224	0.98488	1.00000
VIF	1.84688	14.22680	59.62253	24112.56	1078.89	143136.56

Table 34: Summary statistics of robustness checks for the FFC benchmark

Test	Min	Q1	Median	Mean	Q3	Max
Shapiro-Wilk	0.00000	0.12954	0.32846	0.39014	0.54933	0.99773
Breusch-Pagan	0.00000	0.05212	0.21216	0.32087	0.43031	0.99910
Durbin-Watson	0.00055	0.55580	0.74770	0.65988	0.91555	1.00000
VIF	1.00020	1.08097	1.22805	1.40293	1.30240	2.60217

Table 35: Summary statistics of robustness checks for the CAPM benchmark

Test	Min	Q1	Median	Mean	Q3	Max
Shapiro-Wilk	0e+00	0.06881	0.23795	0.33417	0.44763	0.9992
Breusch-Pagan	NA	NA	NA	NA	NA	NA
Durbin-Watson	7e-05	0.52319	0.69896	0.63190	0.87136	1.0000
VIF	NA	NA	NA	NA	NA	NA

The tables above report summary statistics of robustness checks on the three benchmarks for the Shapiro-Wilk, Breusch-Pagan, Durbin-Watson, and Variance Inflation Factor (VIF) tests. All test reports p-values except for VIF.

The Vanguard benchmark exhibits multicollinearity due to an extremely high VIF, indicating that the coefficients can be distorted and that the t-test and confidence interval in the model is inflated. The expectation in terms of multicollinearity is, however, that as funds invest in the stock market and stocks are highly correlated with each other, that the independent variables also are highly correlated.

7.4 Robustness check using Swedish benchmarks

Table 36: Swedish (Storebrand) benchmark index funds

Name	ISIN	Global Category
Aktiespararna Topp Sverige Hållbar A	SE0000924649	Europe Equity Large Cap
Case Sverige Högutdelande Bolag A	SE0000577272	Europe Equity Large Cap
Handelsbanken Europa Index Crit (A1 SEK)	SE0000624421	Europe Equity Large Cap
KPA Blandfond	SE0000885089	Aggressive Allocation
Länsförsäkringar Japan Index	SE0000837296	Japan Equity
Nordea Inst Aktie Sverige	SE0000524407	Europe Equity Large Cap
Öhman Marknad Europa A	SE0000577975	Europe Equity Large Cap
Öhman Marknad Japan A	SE0000577959	Japan Equity
Öhman Marknad Pacific A	SE0000577983	Asia ex-Japan Equity
Öhman Marknad USA A	SE0000577967	US Equity Large Cap Blend
Skandia Europa Exponering	SE0000810772	Europe Equity Large Cap
Skandia Japan Exponering	SE0000810764	Japan Equity
*Storebrand Europa A SEK	SE0000531881	Europe Equity Large Cap
*Storebrand Global All Countries A SEK	SE0000671919	Global Equity Large Cap
*Storebrand Japan A SEK	SE0000621393	Japan Equity
*Storebrand Sverige A SEK	SE0000529992	Europe Equity Large Cap
*Storebrand USA A SEK	SE0000594111	US Equity Large Cap Blend
Swedbank Robur Access Mix A	SE0000434359	Moderate Allocation

This table reports the list of Swedish index funds obtained in the screening from Morningstar Direct. * denotes that the fund is used in the construction of the Swedish index funds benchmark.

Table 37: Summary statistics of estimated value added ($\hat{S}_i$) for the Swedish benchmarks

Metric	Storebrand	FFC	CAPM
Mean	4 693 208	4 590 608	4 002 031
Standard Deviation	12 457 624	12 907 393	11 782 864
T-Statistic	10.81	10.21	9.75
1st Percentile	-10 638 067	-14 596 926	-15 674 992
5th Percentile	-3 079 762	-5 535 223	-6 202 095
10th Percentile	-1 474 614	-2 002 596	-2 440 829
Median	448 628	487 794	317 887
90th Percentile	14 793 992	16 860 071	15 157 932
95th Percentile	28 003 132	27 924 388	28 293 933
99th Percentile	75 007 884	77 249 499	65 417 446

This table reports descriptive statistics for estimated value added ($\hat{S}_i$) for a fund which exists for T_i periods using the Swedish set of index funds benchmark, the FFC benchmark with factors for the Swedish market and the CAPM benchmark with factors for the Swedish market. Note that the period for the estimation is limited to 2001-2019 due to data availability.

Table 38: Two-sided T-test of the no-skill hypothesis using the Swedish benchmarks

Benchmark	Estimate	Statistic	P_value	DF	Conf_low	Conf_high
Storebrand	4 693 208***	10.81	1.5022e-25	822	3 840 848	5 545 569
FFC	4 590 608***	10.2	4.2683e-23	822	3 707 474	5 473 742
CAPM	4 002 031***	9.74	2.6433e-21	822	3 195 838	4 808 223

This table reports the results from a two-sided t-test for estimated value added ($\hat{S}_i$) using the Swedish set of index funds (Storebrand) benchmark, the FFC benchmark with factors for the Swedish market and the CAPM benchmark with factors for the Swedish market.